

Kazalo / Table of Contents

Predgovor	i
Preface	ii
Organizacijski odbor / Organising committee	iii
Programski odbor / Programme committee	iii
Vodje tematskih področij / Area chairs	iv
Člani programskega odbora / Programme committee members	v
Organizatorji / Organizers	vii
Sponzorji / Sponsors	viii
Urniki / Timetable	ix
Kazalo / Table of Contents	xiv
 INVITED TALKS	 1
Contrasting post-editing and human translation along the dimension of term and cognate variation <i>Oliver Čulo</i>	 1
Speech Is Golden – On ASR at the Service of the Danish Public Sector <i>Peter Juel Henriksen</i>	 4
The Text Encoding Initiative: 30 Years of Accumulated Wisdom and Its Potential for a Bright Future <i>Laurent Romary</i>	 5
The Landscape of Digital Annotation and Its Meaning <i>Niels-Oliver Walkowski</i>	 6
Universal Dependencies for Slavic Languages <i>Daniel Zeman</i>	 12
 PAPERS	 13
Asistent – A Machine Translation System for Slovene, Serbian and Croatian <i>Mihael Arčan, Maja Popović, Paul Buitelaar</i>	 13
Ohranjanje jezikovne zahtevnosti besedil pri prevajanju testov PISA <i>Špela Arhar Holdt, Iztok Kosem</i>	 21
Predstavitveni portal spletnih jezikovnih virov za slovenščino <i>Špela Arhar Holdt, Kaja Dobrovoljc, Iztok Kosem</i>	 27
Integrating Natural Language and Formal Analysis for Legal Documents <i>Shaun Azzopardi, Albert Gatt, Gordon J. Pace</i>	 32
Organiziranje projekta vizualizacije podatkov <i>Narvika Bovcon, Jure Demšar, Aleš Vaupotič</i>	 36
Razvoj učne množice za izboljšano označevanje spletnih besedil <i>Jaka Čibej, Špela Arhar Holdt, Tomaž Erjavec, Darja Fišer</i>	 40

Preizkus Googlevega govornega programskega vmesnika pri samodejnem razpoznavanju govorne slovenščine <i>Simon Dobrišek, David Čefarin, Vitomir Štruc, France Mihelič</i>	47
Povezljivost pravopisnih pravil in slovarja: sanje pravopiscev 20. stoletja <i>Helena Dobrovoljc</i>	52
Slovenska akademska besedila: prototipni korpus in načrt analiz <i>Tomaž Erjavec, Darja Fišer, Nikola Ljubešić, Nataša Logar, Milan Ojsteršek</i>	58
Sentiment Annotation of Slovene User-Generated Content <i>Darja Fišer, Jasmina Smailović, Tomaž Erjavec, Igor Mozetič, Miha Grčar</i>	65
Slovar tviterščine <i>Polona Gantar, Iza Škrjanec, Darja Fišer, Tomaž Erjavec</i>	71
Analiza krajšanja slovenskih sporočil na družbenem omrežju Twitter <i>Teja Goli, Eneja Osrajnik, Darja Fišer</i>	77
Analiza mnenj s pomočjo strojnega učenja in slovenskega leksikona sentimenta <i>Klemen Kadunc, Marko Robnik-Šikonja</i>	83
Building a Gold Standard for Temporal Entity Extraction from Medieval German Texts <i>Natalia Korchagina</i>	90
Šolar 2.0: nadgradnja korpusa šolskih pisnih izdelkov <i>Iztok Kosem, Tadeja Rozman, Špela Arhar Holdt, Polonca Kocjančič, Cyprian Laskowski</i>	95
Baza kolokacijskega slovarja slovenskega jezika <i>Simon Krek, Polona Gantar, Iztok Kosem, Vojko Gorjanc, Cyprian Laskowski</i>	101
Označevanje udeleženskih vlog v učnem korpusu za slovenščino <i>Simon Krek, Polona Gantar, Kaja Dobrovoljc, Iza Škrjanec</i>	106
Priprava usporedivih korpusa za usporedbu <i>Ivana Lalli Pačelat</i>	111
Easily Accessible Language Technologies for Slovene, Croatian and Serbian <i>Nikola Ljubešić, Tomaž Erjavec, Darja Fišer, Tanja Samardžić, Maja Miličević, Filip Klubička, Filip Petkovski</i>	120
Strojno podprta kolacija slovenskih rokopisnih besedil: variantna mesta v luči računalniških algoritmov in vizualizacij <i>Matija Ogrin, Andrejka Žejn</i>	125
Popisi prebivalstva Slovenije 1830–1931: Orodje za transkribiranje historičnih demografskih podatkov <i>Andrej Pančur</i>	133
Označevanje zbirke zapisnikov sej slovenskega parlamenta s smernicami TEI <i>Andrej Pančur</i>	142
Raba vejice v uporabniških spletnih vsebinah <i>Damjan Popič, Darja Fišer, Katja Zupan, Polona Logar</i>	149

EAGLE – Medomrežje Europeana antične grške in latinske epigrafike.	
Digitalni dostop do antičnih napisnih spomenikov	
<i>Anja Ragolič</i>	154
ARIADNE: povezani odprti podatki (LOD) v praksi	
<i>Benjamin Štular, Franco Niccolucci, Julian Richards</i>	158
Digital Video in Digital Humanities Methodology: A Case Study	
<i>Aleš Vaupotič, Marco Buziol, Narvika Bovcon</i>	164
Digitalna humanistika v šoli	
<i>Maja Vičič Krabonja</i>	170
Generiranje kritičnih prepisov s strojnim prevajanjem na ravni znakov	
<i>Katja Zupan, Tomaž Erjavec</i>	175
Sintetizator govora za slovenščino eBralec	
<i>Jerneja Žganec Gros, Boštjan Vesnicher, Simon Rozman, Peter Holozan, Tomaž Šef</i>	180
Razpoznavanje tekočega govora v slovenščini z bazo predavanj SI TEDx-UM	
<i>Andrej Žgank, Darinka Verdonik, Mirjam Sepesy Maučec</i>	186
ABSTRACTS	190
Pretvorba korpusa ssj500k v Univerzalno odvisnostno drevesnico za slovenščino	
<i>Kaja Dobrovoljc, Tomaž Erjavec, Simon Krek</i>	190
What is critical in digital humanities?	
<i>Mario Hibert</i>	193
Asian Language Teaching and Learning - The Influence of Technology on Students' Skills in SL Classroom	
<i>Marijana Janjić, Sara Librenjak, Kristina Kocijan</i>	196
Zbiva in EWD, spletni orodji za arheološke raziskave	
<i>Bojan Kastelic, Mateja Belak, Andrej Pleterski, Benjamin Štular, Miran Erič</i>	198
Nadgradnja korpusov Gigafida, Kres, ccGigafida in ccKres	
<i>Simon Krek, Polona Gantar, Špela Arhar Holdt, Vojko Gorjanc</i>	200
Digitalizacija in uporaba digitalnega etnomuzikološkega zvočnega gradiva – izkušnje Glasbenonarodopisnega inštituta ZRC SAZU	
<i>Drago Kunej</i>	204
Trirazsežno dokumentiranje v službi varovanja nepremične kulturne dediščine	
<i>Aleš Lazar, Sonja Ifko</i>	205
Analysing Spatial Distribution of Linguistic Variables in Geoencoded Tweets from Croatia, Bosnia, Montenegro and Serbia	
<i>Nikola Ljubešić, Tanja Samardžić, Maja Miličević</i>	208
The First World War on the Web - The Case of Serbia	
<i>Slobodan Mandić</i>	210
Od znanstvenokritične izdaje do repozitorija rokopisov: Dosežki, možnosti in načrti	
<i>Matija Ogrin, Tomaž Erjavec, Jan Jona Javoršek</i>	214

The Use of Semantic Word Classes in Document Classification <i>Stevan Ostrogonac, Branislav Popović, Milan Sečujski</i>	216
Dolgotrajno ohranjanje raziskovalnih podatkov v manjših raziskovalnih infrastrukturah: Uporaba odprtokodne aplikacije Archivematica <i>Andrej Pančur, Bogomir Rožman</i>	218
Razvoj aplikacije za spodbujanje trajnostne mobilnosti <i>Dan Podjed, Saša Babič, Tatiana Bajuk Senčar, Alenka Bezjak Mlakar, Gregor Burger, Jurij Fikfak, Jože Guna, Marko Maver, Matevž Pogačnik, Emilija Stojmenova Duh, Uroš Žolnir</i>	220
Luščenje in jezikoslovna analiza kolokacij iz korpusa Šolar <i>Tadeja Rozman, Špela Arhar Holdt, Senja Pollak, Iztok Kosem</i>	222
Historična topografija Slovenije <i>Miha Seručnik</i>	225
Language Technologies in Humanities: Computational Semantic Analysis in Folkloristics <i>Gregor Strle, Matija Marolt</i>	227
Slikovna retrospektiva porušenega Breginja in analiza pokrajinskih sprememb <i>Tatjana Veljanovski, Žiga Kokalj</i>	230
STUDENT PAPERS	233
Gradnja in analiza petjezičnega korpusa podnaslovov govorov TED <i>Miha Helbl, Žiga Domevšček</i>	233
#Analiza novih komunikacijskih elementov na družbenem omrežju @Twitter <i>Katerina Pertot, Maja Petrovčič, Nika Strojan</i>	239
Razdvoumljanje besednega pomena pri strojnih prevajalnikih Amebis Presis, Google Translate in MT@EC <i>Jure Škerl</i>	245
Korpus študentov prevajanja MetaTrans <i>Anja Tavčar, Ines Čeligoj Pregelj, Miha Pompe</i>	250
STUDENT ABSTRACTS	256
Literature Reloaded: using databases to explore literary trends <i>Lucija Dačić</i>	256
Digitalna arheologija? Primer uporabe digitalnih orodij za analizo arheološkega najdišča <i>Jernej Rihter</i>	258